The Trauma Mortality Prediction Model *(-tmpm.ado-)* is Robust to the AIS, ICD-9, and the ICD-10 Lexicons

Alan Cook, MD, FACS

Turner Osler, MD, MS (Biostatistics), FACS

Stata Conference, San Diego

July 26-27, 2012

- Objectives
  - Briefly Discuss :
    - The importance of mortality prediction in trauma
    - The foregoing mortality prediction models in trauma
  - Introduce the Trauma Mortality Prediction Model (TMPM)
  - Highlight the ~tmpm.ado~ user-written command for Stata

TMPM: Trauma Mortality Prediction Model

TMPM: Trauma Mortality Prediction Model

10 Leading Causes of Death by Age Group, United States – 2009

	Age Groups										
Rank	<1	1-4	5-9	10-14	15-24	25-34	35-44	45-54	55-64	65+	Total
1	Congenital Anomalies 5,319	Unintentional Injury 1,466	Unintentional Injury 773	Unintentional Injury 916	Unintentional Injury 12,458	Unintentional Injury 14,062	Unintentional Injury 15,102	Malignant Neoplasms 50,616	Malignant Neoplasms 106,829	Heart Disease 479,150	Heart Disease 599,413
2	Short Gestation 4,538	Congenital Anomalies 464	Malignant Neoplasms 477	Malignant Neoplasms 419	Homicide 4,862	Suicide 5,320	Malignant Neoplasms 12,519	Heart Disease 36,927	Heart Disease 67,261	Malignant Neoplasms 391,035	Malignant Neoplasms 567,628
3	SIDS 2,226	Homicide 376	Congenital Anomalies 195	Suicide 259	Suicide 4,371	Homicide 4,222	Heart Disease 11,081	Unintentional Injury 19,974	Chronic Low. Respiratory Disease 14,160	Chronic Low Respiratory Disease 117,098	Chronic Low. Respiratory Disease 137,353
4	Maternal Pregnancy Comp. 1,608	Malignant Neoplasms 350	Homicide 119	Homicide 186	Malignant Neoplasms 1,636	Malignant Neoplasms 3,659	Suicide 6,677	Suicide 8,598	Unintentional Injury 12,933	Cerebro-vascular 109,238	Cerebro-vascular 128,842
5	Unintentional Injury 1,181	Heart Disease 154	Influenza & Pneumonia 106	Congenital Anomalies 169	Heart Disease 1,035	Heart Disease 3,174	Homicide 2,762	Liver Disease 8,377	Diabetes Mellitus 11,361	Alzheimer's Disease 78,168	Unintentional Injury 118,021
6	Placenta Cord. Membranes 1,064	Influenza & Pneumonia 146	Heart Disease 97	Influenza & Pneumonia 122	Congenital Anomalies 457	HIV 881	Liver Disease 2,481	Cerebro-vascular 6,163	Cerebro-vascular 10,523	Diabetes Mellitus 48,944	Alzheimer's Disease 79,003
7	Bacterial Sepsis 652	Septicemia 71	Chronic Low. Respiratory Disease 64	Heart Disease 120	Influenza & Pneumonia 418	Influenza & Pneumonia 807	HIV 2,425	Diabetes Mellitus 5,725	Liver Disease 9,154	Influenza & Pneumonia 43,469	Diabetes Mellitus 68,705
8	Respiratory Distress 595	Chronic Low. Respiratory Disease 66	Benign Neoplasms 40	Chronic Low. Respiratory Disease 59	Complicated Pregnancy 227	Diabetes Mellitus 604	Cerebro-vascular 1,916	Chronic Low. Respiratory Disease 4,664	Suicide 5,808	Nephritis 40,465	Influenza & Pneumonia 53,692
9	Circulatory System Disease 581	Perinatal Period 58	Septicemia 33	Benign Neoplasms 45	Cerebro-vascular 193	Cerebro-vascular 537	Diabetes Mellitus 1,872	HIV 3,388	Nephritis 4,792	Unintentional Injury 39,111	Nephritis 48,935
10	Neonatal Hemorrhage 517	Benign Neoplasms 53	Cerebro-vascular 32	Cerebro-vascular 42	Chronic Low. Respiratory Disease 187	Liver Disease 459	Influenza & Pneumonia 1,314	Influenza & Pneumonia 2,918	Septicemia 4,628	Septicemia 26,763	Suicide 36,909

Data Source: National Vital Statistics System, National Center for Health Statistics, CDC.
Produced by: Office of Statistics and Programming, National Center for Injury Prevention and Control, CDC using WISQARS™.

Centers for Disease Control and Prevention
National Center for Injury Prevention and Control

CS230409

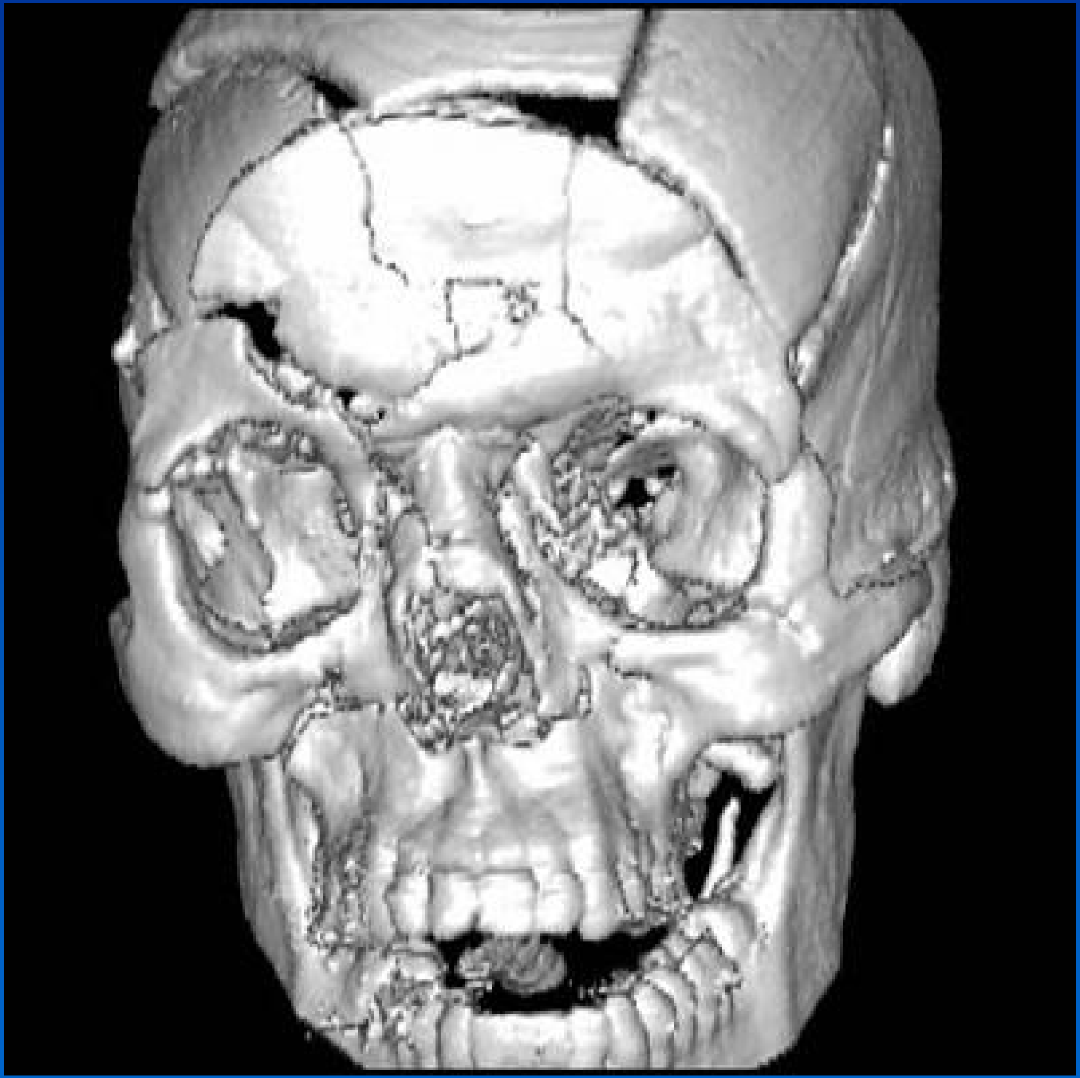
TMPM: Trauma Mortality Prediction Model

- Why estimate mortality in populations of injured patients?
  - The quantification of trauma is required for:
    - The scientific study of injury
    - The objective evaluation of trauma care
  - We compromise our research and our patient care if we fail to develop and use accurate measures of trauma outcomes
    - Specifically mortality

TMPM: Trauma Mortality Prediction Model

- If a human being goes through a windshield at 80 mph, it might seem that there are an infinite number of different possible injuries could result.
- A different spectrum of injuries could result from a stab wound or gunshot wound to the chest or abdomen

TMPM: Trauma Mortality Prediction Model



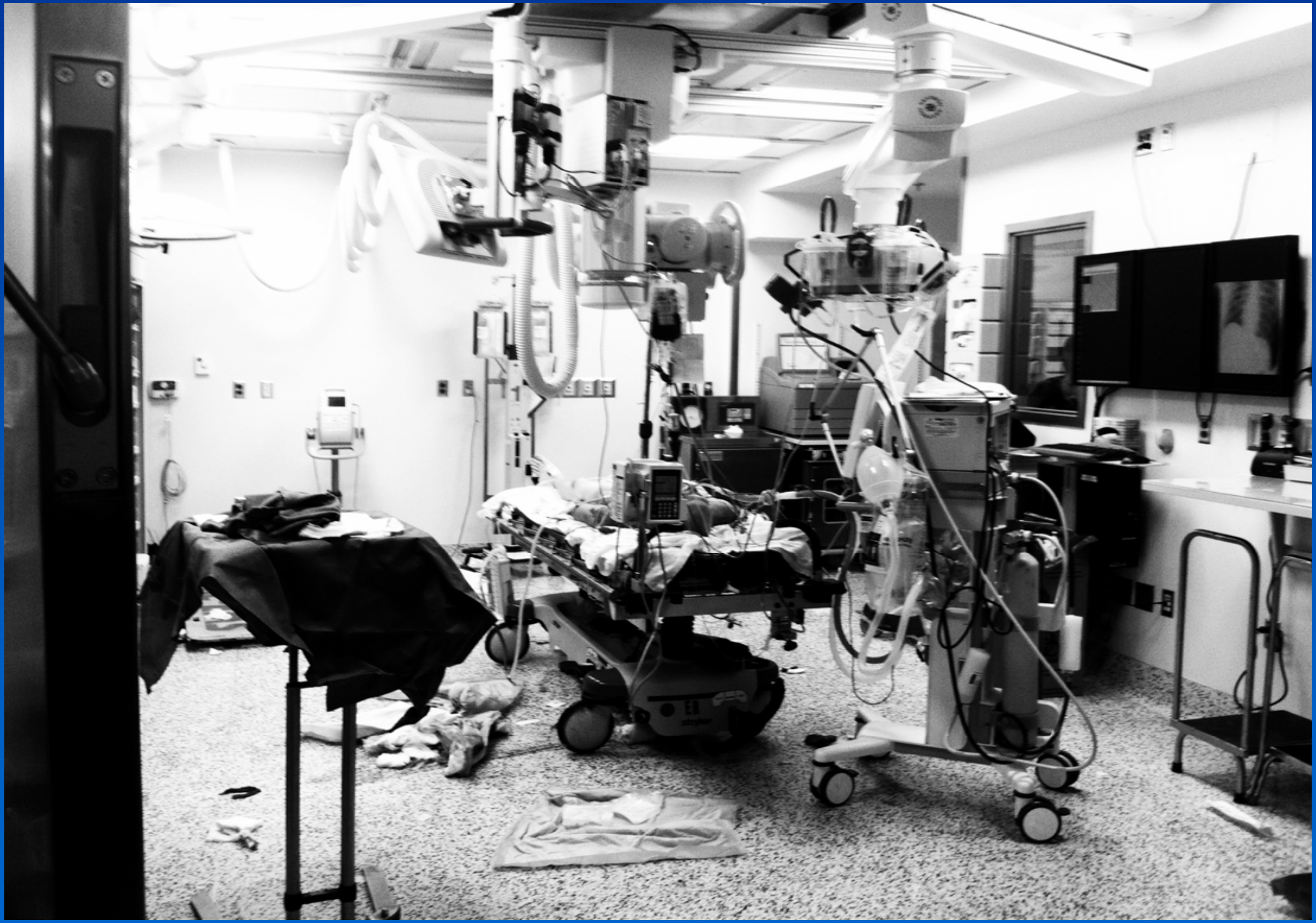
TMPM: Trauma Mortality Prediction Model

- Human beings are structurally similar to one another
  - Human anatomy tends to “fail” in structurally similar ways
  - There are only ~1,000 different injuries a person could have
    - Fortunately, these injuries have already been extravagantly detailed in injury coding lexicons such as the AIS and ICD-9 coding systems

TMPM: Trauma Mortality Prediction Model

- Unfortunately, a person could have up to, say 10 injuries
- So the number of injury *patterns* one could have is 10^1000
  - This is more than the number of atoms in the universe: 10^80

TMPM: Trauma Mortality Prediction Model



TMPM: Trauma Mortality Prediction Model

- Mortality estimation is not new in the field of trauma surgery
  - Injury Severity Score (ISS)
    - Based on the Abbreviated Injury Score
    - Baker, S., B. O'Niell, et al. (1974). "The injury severity score:A method for describing patients with multiple injuries and evaluating emergency care." J Trauma **14**(3): 187-196.
  - Trauma Score & ISS (TRISS)
    - Boyd, C., M. Tolson, et al. (1987). "Evaluating trauma care:The TRISS method." J Trauma **27**(1): 370-378.
  - International Classification of Disease-9 (ICD-9) based injury severity score (ICISS)
    - Osler, T., R. Rutledge, et al. (1996). "ICISS: an international classification of disease-9 based injury severity score." J Trauma **41**(3): 380-386; discussion 386-388

TMPM: Trauma Mortality Prediction Model

- Most mortality prediction models are based on two administrative coding systems
  - International Classification of Diseases, 9th Revision, Clinical Modification, (ICD-9 CM or ICD-9)
    - Primarily a taxonomy of ALL human disease, including traumatic injuries
    - The basis of billing transactions between providers and payers
  - Abbreviated Injury Score
    - A taxonomy of traumatic injury and includes a severity score
    - Used in trauma data registries and mortality risk models

TMPM: Trauma Mortality Prediction Model

ICD-9 Lexicon for Coding Liver Injuries

✓4th **864 Injury to liver**

The following fifth-digit subclassification is for use with category 864:

- **0 unspecified injury**
- **1 hematoma and contusion**
- **2 laceration, minor**
  - Laceration involving capsule only, or without significant involvement of hepatic parenchyma [i.e., less than 1 cm deep]
- **3 laceration, moderate**
  - Laceration involving parenchyma but without major disruption of parenchyma [i.e., less than 10 cm long and less than 3 cm deep]
- **4 laceration, major**
  - Laceration with significant disruption of hepatic parenchyma [i.e., 10 cm long and 3 cm deep]
  - Multiple moderate lacerations, with or without hematoma
  - Stellate lacerations of liver
- **5 laceration, unspecified**
- **9 other**

TMPM: Trauma Mortality Prediction Model

AIS Lexicon for Coding of Liver Injuries

541899.2	Liver NFS
541810.2	contusion (hematoma) NFS
541812.2	subcapsular, ≤ 50% surface area, nonexpanding or intraparenchymal ≤ 10cm in diameter; minor; superficial; (OIS Grade I or II)
541814.3	> 50% surface area or expanding; ruptured subcapsular or parenchymal; intraparenchymal > 10cm or expanding; blood loss > 20% by volume; major; subcapsular; (OIS Grade III)
541820.2	laceration NFS
541822.2	simple capsular tears, ≤ 3 cm parenchymal depth, ≤10cm in length; blood loss ≤ 20% by volume; minor; superficial (OIS Grade I or II).
541824.3	> 3 cm parenchymal depth ; major duct involvement"; blood loss > 20% by volume; moderate (OIS Grade III)
541826.4	parenchymal disruption of ≤ 75% of hepatic lobe or 1-3 Couinaud's segments within a single lobe; multiple lacerations > 3 cm deep; "burst" injury; major (OIS Grade IV)
541828.5	parenchymal disruption of > 75% of hepatic lobe or involving > 3 Couinard's segments within a single lobe or involving retrohepatic vena cava/central hepatic veins; massive; complex (OIS Grade V)
541830.6	hepatic avulsion (total separation of all vascular attachments) (OIS Grade VI)
641840.4	rupture ("fracture") NFS Use this code only when a more detailed description is not available.

TMPM: Trauma Mortality Prediction Model

- A third administrative coding system is on the horizon for US healthcare
  - ICD-10
    - Industry standard for most of the world
    - Like the metric system, not implemented in the US yet
      - US Department of Health and Human services delayed compliance to October 1, 2013, then delayed again to October 2014

TMPM: Trauma Mortality Prediction Model

ICD-10 Lexicon for Coding Liver Injuries

✓6th **S36.11 Injury of liver**

✓7th **S36.112 Contusion of liver**

✓7th **S36.113 Laceration of liver, unspecified degree**

✓7th **S36.114 Minor laceration of liver**

Laceration involving capsule only, or, without significant involvement of hepatic parenchyma [i.e., less than 1 cm deep]

✓7th **S36.115 Moderate laceration of liver**

Laceration involving parenchyma but without major disruption of parenchyma [i.e., less than 10 cm long and less than 3 cm deep]

✓7th **S36.116 Major laceration of liver**

Laceration with significant disruption of hepatic parenchyma [i.e., greater than 10 cm long and 3 cm deep]

Multiple moderate lacerations, with or without hematoma

Stellate laceration of liver

✓7th **S36.118 Other injury of liver**

✓7th **S36.119 Unspecified injury of liver**

TMPM: Trauma Mortality Prediction Model

- <u>The Trauma Mortality Prediction Model</u> (TMPM)
  - We wanted to be able to provide an estimated probability of mortality for any of these 10^1000 injury patterns
  - How are we to proceed?
  - Given a large enough dataset, one might estimate the probability of producing death for each of the 1,000 possible injuries
    - We'll call that "severity" or the "<u>dose of trauma"</u>

TMPM: Trauma Mortality Prediction Model

- TMPM Development
  - The basic idea was to fit a probit model with death as the outcome with 1,000 possible injuries as binary predictors
    - The cohort was the National Trauma Data Bank (NTDB)
      - N=1,000,000
    - Design matrix was 1,000 variables wide and 1,000,000 cases long
      - Mostly zeros – Most patients don't have most injuries
      - The sum of the dataset provided a rich mosaic of injuries with the associated patient survival

- TMPM Development
  - Two separate probit models were created using the injuries described in the coding system
    - Model 1 used all possible injuries as binary predictors and death as a binary outcome
    - Model 2 was based on body region severity indicators
  - The empiric injury severity for each injury was estimated by taking a weighted average of the coefficients of the 2 regression models

TMPM: Trauma Mortality Prediction Model

- Why two probit models?
  - Even in the NTDB dataset, the prevalence of specific injuries was very uneven.
    - A few injuries were very common (minor skin laceration of the face)
    - Many injuries were very rare
      - Twenty-four (2%) AIS codes occurred only once in the entire dataset.
      - 684 codes (52%) occurred fewer than 100 times
  - To estimate the severity of these rare injuries, they were “lumped” into larger groups
    - Now we have 2 estimates for each injury – so now what?
    - Average the 2 models weighted on the inverse of the variance of each of the 2 estimates!

TMPM: Trauma Mortality Prediction Model

- MARC: Model Averaged Regression Coefficients
  - How the models were averaged

$$\frac{\frac{1}{Var(Model\,1)}}{\frac{1}{Var(Model\,1)}+\frac{1}{Var(Model\,2)}}Coef_{Model\,1}+\frac{\frac{1}{Var(Model\,2}}{\frac{1}{Var(Model\,1)}+\frac{1}{Var(Model\,2)}}Coef_{Mode\,2}$$

$$Var(RScoef_m)=\frac{1}{N_i-1}\sum_{j=1}^{N_i}N_i\times W_j\times\left[\left(AIScoef_j\right)-E\left(AIScoef\right)\right]^2$$

Where:

N_i is the number of AIS codes in RS region m

W_j is the weight of the contribution of $AIScoef_j$ to the variance of $RScoef_m$

$$W_j=\frac{\frac{1}{Var\ AIScoef_j}}{\sum_{r=1}^{N_i}\frac{1}{Var\ AIScoef_r}}$$

$$E(AIScoef)=\sum_{j=1}^{N_i}W_j\times AIScoef_j$$

TMPM: Trauma Mortality Prediction Model

- Ultimately the probability of death or p(Death) is computed as follows:

$$p(Death) = \Phi\left[\begin{array}{l} C_0 + C_1 \times I_1 + C_2 \times I_2 + C_3 \times I_3 + C_4 \times I_4 + \\ C_5 \times I_5 + C_6 \times S + C_7 \times I_1 \times I_2 \end{array}\right]$$

Where Φ is the inverse normal function

C_x are the coefficients from the model(s)

I_x are the marc values of the 5 worst injuries

S is an indicator variable indicating whether the worst 2 injuries are in the same body region

TMPM: Trauma Mortality Prediction Model

TMPM vs. Other ICD-9 Based Mortality Prediction Scores

Model	ROC	HL Stat	AIC
ICISS	0.846	432	44,071
SWI: SRR	0.861	833	42,813
SWI: MARC	0.872	112	42,716
All Injury: MARC	0.878	357	42,688
TMPM-ICD9	0.880	19	41,251

TMPM: Trauma Mortality Prediction Model

TMPM vs. Other AIS Based Mortality Prediction Scores

Model	ROC**	HL Stat*	AIC*
ISS	0.872	296	37225
SWI	0.891	314	34059
TMPM	0.902	58	32003

Model	ROC**	HL Stat*	AIC*
ISS age+gender+mech	0.915	54	33773
SWI age+gender+mech	0.921	128	31770
TMPM age+gender+ mech	0.928	19	29645

TMPM: Trauma Mortality Prediction Model

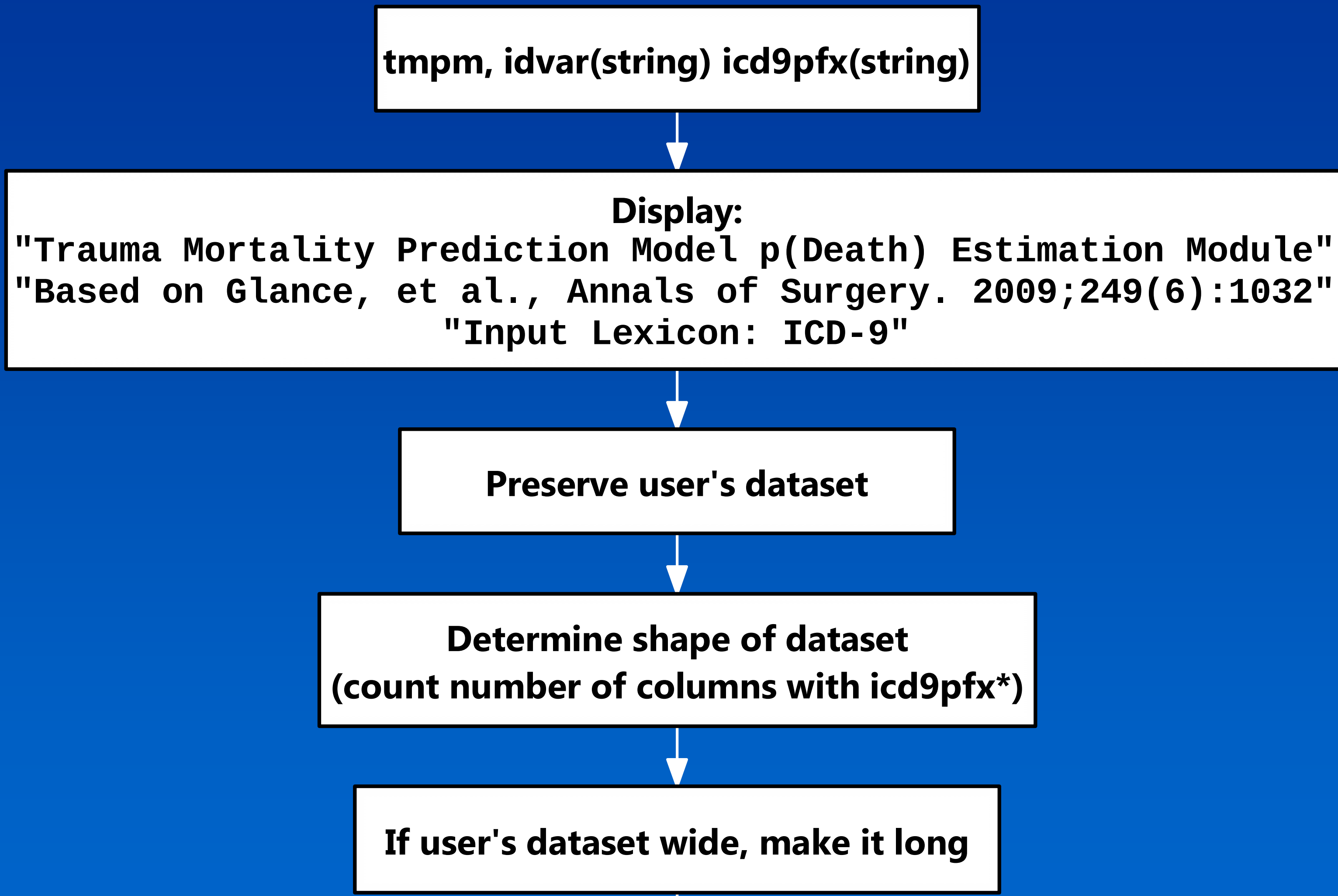
TMPM: Trauma Mortality Prediction Model

- We used Stata to understand the problem and create a prediction model.
- Unfortunately, the prediction model was complicated enough that we feared users might not be able to implement it easily and accurately.
- So, we used Stata in a second context, to create a simple command that would compute our model correctly, no matter who was using our model...."

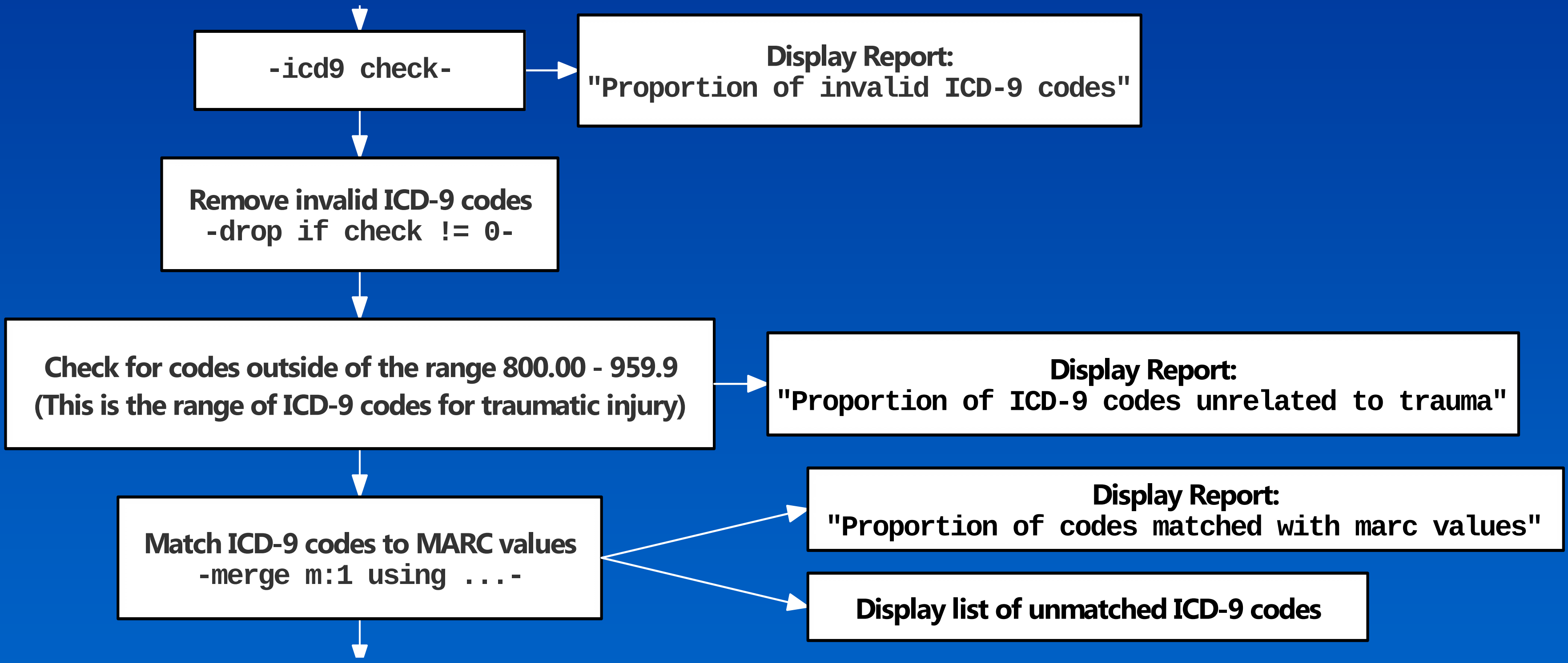
- ~tmpm.ado~
  - Goals
    - Execute the TMPM calculation of p(Death) correctly
    - Design ~tmpm.ado~ to accommodate AIS, ICD-9 or ICD-10 lexicons from a single command line
      - One at a time, though
    - Streamline the command itself
      - Can recognize if the data are formatted as wide or long without the user specifying such in the command
      - Give the user the option of receiving information about their dataset while the program is running

- ~tmpm.ado~
  - Three modules to the program
    - ICD-9
    - AIS
    - ICD-10
  - `syntax [varlist], [ idvar(varname) aispfx(string) icd9pfx(string) icd10pfx(string) NOREPORT ]`
  - The appropriate module is called based on which ~pfx is identified in the command line
  - The option "NOREPORT" suppresses the output of reports about the user's dataset

TMPM: Trauma Mortality Prediction Model



TMPM: Trauma Mortality Prediction Model



TMPM: Trauma Mortality Prediction Model

Keep two worst injuries and identify if they are from same body region (this is a variable from the MARC table)

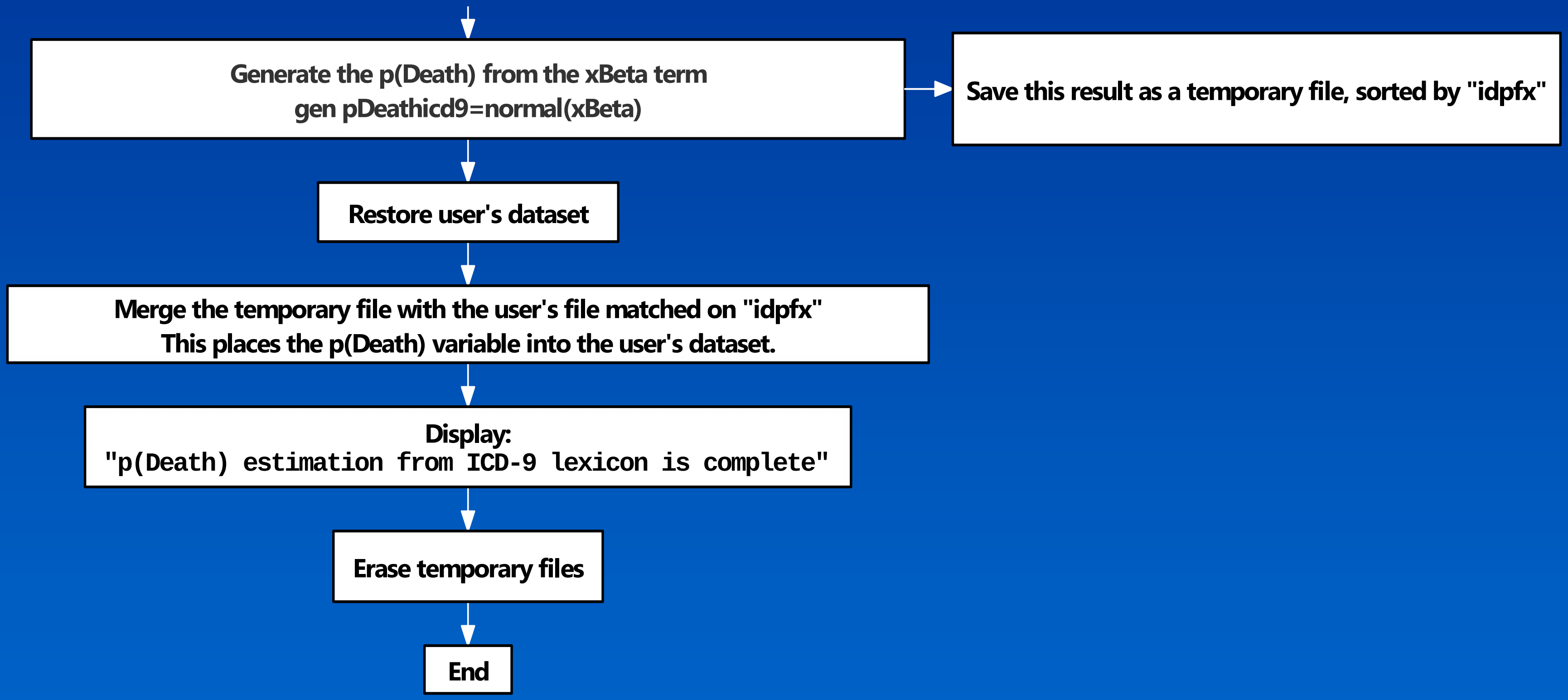
Keep five worst (highest) marc values

Generate interaction term for two worst injuries

Calculate the model ("xBeta")

```
gen xBeta =  (1.406958*marc1) +  ///
       (1.409992* marc2)     + ///
       (0.5205343* marc3)    + ///
       (0.4150946* marc4)    + ///
       (0.8883929* marc5)    + ///
    (-(0.0890527)* same_region)  + ///
      (-(0.7782696)* Imarc)    + ///
              -(2.217565)
```

TMPM: Trauma Mortality Prediction Model



TMPM: Trauma Mortality Prediction Model

~tmpm.ado output from data using the AIS injury coding system

```
. tmpm, idvar(id) aispfx(predot_)
______________________________________________________________
Trauma Mortality Prediction Model p(Death) Estimation Module
Based on Osler, et al., Annals of Surgery. 2008;247(6):1041
Input Lexicon: AIS
______________________________________________________________

Proportion of codes matched with marc values:        98.75%

Your data contain 434 unique AIS codes

Your unmatched AIS codes:

  Unmatched |
  AIS Codes |      Freq.     Percent        Cum.
------------+-----------------------------------
     115099 |         42       65.63       65.63
     115299 |          5        7.81       73.44
     131806 |          1        1.56       75.00
     515099 |          1        1.56       76.56
     919602 |          3        4.69       81.25
     999999 |         12       18.75      100.00
------------+-----------------------------------
      Total |         64      100.00

p(Death) estimation from AIS lexicon is complete
```

TMPM: Trauma Mortality Prediction Model

```
. tmpm, idvar(id) icd9pfx(icdinj)

Trauma Mortality Prediction Model p(Death) Estimation Module
Based on Glance, et al., Annals of Surgery. 2009;249(6):1032
Input Lexicon: ICD-9


Proportion of invalid ICD-9 codes:          1.07%

Proportion of ICD-9 codes unrelated to trauma:          9.64%
Proportion of codes matched with marc values:         98.57%

Your data contain 543 unique ICD-9 codes

Your unmatched ICD-9 codes:

  Unmatched |
ICD-9 Codes |      Freq.     Percent        Cum.
------------+-----------------------------------
      905.4 |          1        2.04        2.04
      920.0 |         38       77.55       79.59
      930.9 |          1        2.04       81.63
     941.28 |          1        2.04       83.67
      946.0 |          1        2.04       85.71
     948.00 |          1        2.04       87.76
      958.4 |          3        6.12       93.88
      958.8 |          2        4.08       97.96
     959.11 |          1        2.04      100.00
------------+-----------------------------------
      Total |         49      100.00

p(Death) estimation from ICD-9 lexicon is complete
```

TMPM: Trauma Mortality Prediction Model

- ~tmpm.ado~ is structured to accommodate the ICD-10 lexicon
  - Currently it uses a translation from ICD-9 marc values
    - A warning of this appears in the Results window in place of the journal citation
    - This warning is made because, as yet, there are no ICD-10 data to test our predictions against.
    - Since our model is untested, we suggest that it only be used for research purposes until we have some idea of how well it works.

TMPM: Trauma Mortality Prediction Model

~tmpm.ado output from data using the ICD-10 injury coding system

```
. tmpm, idvar(id) icd10pfx(icd10) noreport
________________________________________________________________
Trauma Mortality Prediction Model p(Death) Estimation Module

WARNING: The TMPM ICD-10 module is based on the NIH/CDC mapping
algorithm and has not been evaluated empirically. The authors
advise that it not be used for actual risk stratification.

Input Lexicon: ICD-10
________________________________________________________________

p(Death) estimation from ICD-10 lexicon is complete
```

- Future Directions
  - ~tmpm.ado~ works on Stata 11 and 12
    - We are seeking to retrofit it to earlier versions of Stata.
  - Implement a (click to run) feature in the Examples portion of the help file

Thank you

The Trauma Mortality Prediction Model

Thanks to
Kreshna Gopal, Ph.D.
Senior Computer Scientist, StataCorp

Nicholas Cox, Ph.D.
Department of Geography, Durham University
Editor, The Stata Journal
Frequent contributor to Statalist

BAYLOR University Medical Center at Dallas
Level I Trauma Center

Stata Data Analysis and Statistical Software

The University of Vermont College of Medicine